

ÉTUDE SUR LE POUVOIR DE PRÉVISION DE L'INDICE D'OFFRE D'EMPLOI

Thierno A. Baldé, Norma Chhab-Alperin and Benoît Quenneville¹

RÉSUMÉ

Cette étude porte sur le pouvoir de prévision de l'indice d'offre d'emploi (IOE) par rapport à quelques variables macroéconomiques telles l'emploi et le chômage. A travers une analyse statistique de causalité faisant intervenir les modèles de séries chronologiques, nous montrons que la relation (ou corrélation) qui semble exister entre l'IOE et les variables d'intérêt n'est pas nécessairement de nature causale. En plus, nous montrerons qu'il est difficile d'obtenir une forme analytique de cette relation qui puisse être utilisable à des fins de prédiction.

MOTS CLÉS: causalité de Granger; fonctions de transfert; lissage; prévision; modèles ARIMA; moyenne mobile.

ABSTRACT

This study concerns the predictive power of the Help Wanted Index (HWI) in comparison to some macroeconomic variables such as Employment and Unemployment. Using a statistical causality analysis involving time series modelling, we show that the relation (or correlation) that seems to exist between HWI and the target variables is not necessarily a causal one. Furthermore, we show that it is not easy to obtain an analytical form of this relation that is usable for prediction purposes.

KEYWORDS: ARIMA models, Forecasting Moving average, Smoothing, Granger causality, Transfer functions models.

1. INTRODUCTION

Cette étude porte sur le pouvoir de prévision de l'indice d'offre d'emploi (IOE) par rapport à quelques variables macroéconomiques telles l'emploi et le chômage. L'indice d'offre d'emploi a été mis au point à partir des annonces d'offres d'emploi faites dans les journaux (dans la rubrique *annonces classés*) par les employeurs dans le but d'attirer des employés potentiels. Plus précisément, il regroupe les annonces classées de 22 journaux majeurs de régions métropolitaines à travers le Canada. Les comptes de l'année 1996 servent de données de référence et sont donc indexés à 100. L'indice est calculé une fois par mois et publié dans la première (ou deuxième semaine) du mois suivant la compilation.

L'IOE a été introduit en 1973 par le ministère des finances. Il a été adopté par Statistique Canada en 1974 et il était présenté jusqu'à récemment² comme un indicateur avancé sur les conditions du marché de l'emploi. Plusieurs études antérieures conduites à Statistique Canada (voir Haggard-Guénette 1989, Amoah 2000) et dans quelques institutions financières ont montré que la performance de l'indice était satisfaisante, même si la quantification de son avance par rapport aux autres variables pertinentes du marché de l'emploi n'était pas toujours effective.

L'avènement de l'Internet et son utilisation, à la fois comme "medium" de recrutement par les employeurs et comme outil de recherche de travail par les chercheurs d'emploi, ont conduit les analystes à se questionner sur la pertinence de l'indice d'offre d'emploi en tant qu'indicateur avancé des conditions du marché de l'emploi. Aussi, on reproche souvent à l'IOE de sous-estimer la demande d'emploi étant donné qu'il est compilé à partir seulement de la rubrique des annonces classées, excluant d'office les offres d'emploi de la section carrière des journaux, les postes offerts dans la fonction publique et les emplois obtenus via les réseaux sociaux (bouche à oreille etc.). Dans une étude récente consacrée à l'impact de l'Internet sur la performance de l'IOE, Akyeampong et Lacroix (2002) soutiennent qu'en l'état actuel des choses, il n'y a pas suffisamment d'évidence pour affirmer que l'Internet réduit, de manière significative, le pouvoir de prévision de l'IOE. De plus, ces auteurs ajoutent que même si une détérioration de l'indice arrivait à être détectée, elle ne serait pas attribuable au

¹ Thierno A. Baldé, Norma Chhab-Alperin and Benoît Quenneville, Centre de recherche et d'analyse en séries chronologiques, Statistique Canada, Ottawa, Ontario, K1A 0T6, Canada.

² L'IOE a cessé d'être publié depuis le mois d'avril 2003.

seul fait de l'Internet et que d'autres facteurs tels l'intensité et la direction de la sous couverture de la demande d'emploi pourraient aussi avoir contribué à cette détérioration.

Dans ce travail, nous tentons d'apporter quelques éléments de réponse sur les interrogations relatives au pouvoir de prévision de l'IOE par rapport à quelques variables macroéconomiques du domaine de l'emploi. La première section consiste en une analyse graphique qui permet d'une part de montrer la corrélation entre l'IOE et les variables cibles et d'autre part, de faire ressortir visuellement l'avance de l'IOE sur ces variables pertinentes. La deuxième section consiste en une analyse statistique de causalité (principalement l'approche de Granger 1969) qui permettra de tester le pouvoir de prévision de l'IOE par rapport aux variables pertinentes. La question fondamentale qui se pose ici consiste à savoir si la prédiction des variables pertinentes à partir de leur passé conjoint avec l'IOE est meilleure que celle obtenue avec leur seul passé. La troisième section consiste en une modélisation univariée de toutes les séries et une modélisation multivariée faisant intervenir l'IOE comme une variable exogène. L'objectif est de comparer les prévisions obtenues par les deux types de modèles. La dernière section conclut l'étude.

2. ANALYSE GRAPHIQUE

Dans cette section nous procédons à une analyse graphique qui permet d'une part de montrer la corrélation entre l'IOE et les variables pertinentes de l'emploi: le taux d'emploi, le niveau d'emploi et le taux de chômage et d'autre part, de faire ressortir visuellement l'avance de l'IOE sur ces variables pertinentes. Le but ultime est de montrer qu'un changement important tel un retournement dans l'évolution de l'IOE sera suivi d'un retournement dans les autres variables pertinentes et inversement, qu'un changement dans ces variables est annoncé à l'avance par l'IOE. La notion de retournement se réfère ici aux sommets et aux creux de la série chronologique considérée.

Dans la pratique, il n'existe pas de critère universel qui permet de déterminer avec exactitude s'il y a retournement dans les tendances observées. Dans ce cas-ci, nous avons choisi, le critère ad hoc *plus ou moins trois mois*. Selon ce critère, une période de croissance ou de décroissance n'est considérée comme telle que si elle persiste (ou se maintient) pendant au moins trois mois consécutifs. Les points de retournements classiques qui ne satisfont pas au critère ad hoc sont appelés dans le texte *faux signaux*.

Il importe de souligner ici que les séries utilisées dans cette analyse graphique ont été lissées à l'aide d'une moyenne mobile de longueur égale à sept mois de telle sorte que la valeur lissée correspondante à une période donnée est la somme pondérée des trois valeurs précédentes, de la valeur courante et des trois valeurs suivantes (d'où l'appellation lissage v-3)³. Le lissage permet de mieux identifier les tendances et déterminer de façon plus réaliste les points de retournements⁴. Nous avons aussi utilisé une moyenne mobile de 51 termes (lissage v-25) pour servir de *base de comparaison* dans l'identification des points de retournement et des faux signaux. Une telle moyenne mobile ne comporte évidemment pas de faux signaux à cause de sa longueur qui permet d'obtenir une série parfaitement lisse.

Les figures 6.1 et 6.2 de l'appendice présentent, pour la période allant de janvier 1981 à avril 2003, les graphiques simultanés de l'IOE avec les séries cibles. Pour ce qui est du niveau d'emploi, la série a une forte tendance linéaire à la hausse qui rend difficile la « superposition » des graphiques. Pour contourner cette difficulté nous avons extrait cette tendance linéaire et considéré pour la suite, la série du niveau d'emploi sans tendance linéaire (NEST). La figure 6.2 présente les graphiques simultanés de l'IOE avec chacune des séries cibles pour le filtre de base v-25 (datation des points de retournement). La table 6.2 présente les points de retournement et l'avance (exprimée en mois) de l'IOE sur les différentes séries (sous l'utilisation du lissage v-25). Globalement, on dénote une avance de l'IOE sur chacune des séries cibles.

L'examen de la figure 6.1 nous montre que l'IOE, le taux d'emploi et le niveau d'emploi, ont la même allure générale (les courbes évoluent ensemble). Les sommets de l'IOE sont fortement reliés aux sommets de ces deux séries, tout en les précédant. De même, les creux de l'IOE sont fortement reliés aux creux des deux séries, tout en les précédant.

Dans le cas du taux d'emploi, la table 6.1 nous montre un total de neuf mauvaises prédictions de l'IOE: six sont des retournements du taux d'emploi non anticipés par l'IOE (identifiés par le signe + dans la table 6.1) et trois sont des

³ v-3 est utilisé en pratique pour lisser l'IOE à des fins de publication. Généralement, les filtres ont été construits de façon empirique. Leur fonction de gain les caractérisent comme étant des filtres « passe-bas ».

⁴ Si on se limitait aux données publiées mensuellement, comme celle du journal *Le Quotidien* par exemple, le critère ad hoc retenu (plus ou moins 3 mois) n'identifieraient presque aucun point de retournement parce que les fluctuations d'un mois à l'autre sont trop nombreuses pour permettre à une tendance de se maintenir.

retournements de l'IOE qui n'ont pas été suivis par des retournements du taux d'emploi (identifiés par le signe -). Aussi, six des huit mauvaises prédictions se sont produites dans les dix dernières années, c'est à dire après le mois de janvier 1992. Ce dernier point tend à confirmer que la performance de l'IOE, par rapport à la prédiction des points de retournements du taux d'emploi, s'est progressivement dégradée au cours des dix dernières années.

Dans le cas du niveau d'emploi, six des huit mauvaises prédictions sont des retournements non anticipés par l'IOE et deux sont des retournements de l'IOE qui n'ont pas été suivis par des retournements du taux d'emploi. En plus, contrairement aux résultats précédents, quatre des huit mauvaises prédictions se sont produites dans les dix premières années, c'est à dire avant le mois de janvier 1992. Ce dernier point ne semble pas confirmer une dégradation progressive de la performance de l'IOE par rapport au niveau d'emploi. Il serait donc difficile d'imputer à l'Internet la dégradation observée dans les dernières années, particulièrement en 2002.

Dans le cas du taux de chômage, la figure 6.1 montre une forte corrélation négative avec l'IOE. Ce résultat demeure conforme aux attentes théoriques. De plus, trois des sept mauvaises prédictions sont des retournements non anticipés par l'IOE et quatre sont des retournements de l'IOE qui n'ont pas été suivis par le taux de chômage (voir figure 6.1).

3. ANALYSE DE CAUSALITÉ

Dans le domaine de la macroéconomie, la question de savoir si une variable économique peut servir à prévoir une autre variable économique revêt une grande importance. L'approche de Granger (1969) et Sims (1972) aborde la causalité en termes de prédiction d'une variable Y à partir de son passé conjoint avec une autre variable X . La question qui se pose est de savoir si la prédiction de Y à partir de son historique conjoint est significativement meilleure à la prédiction faite par rapport à son seul passé.

La méthode utilisée dans ce travail, consiste à régresser la variable Y sur les valeurs de son historique conjoint, c'est à dire, ses valeurs retardées et celles de la variable X . Les tests de Wald sont alors utilisés pour tester si les paramètres associés aux retards de la variable X sont significatifs. Le modèle de régression considéré dans ce cas-ci est de la forme:

$$Y_t = \alpha_0 + \sum_{i=1}^I \beta_i Y_{t-i} + \sum_{j=1}^J \gamma_j X_{t-j} + \epsilon_t \text{ et le test de Wald porte alors sur les paramètres } \gamma_j .$$

L'avantage principal de cette méthode réside dans le fait qu'elle précise la direction de la causalité et permet de tester celle-ci dans les deux sens. L'équation précédente permet de tester si la variable X cause la variable Y à la Granger, mais l'on peut tester aussi la relation symétrique (Y cause X) en utilisant l'équation de régression de la variable X sur ses valeurs retardées et celles de la variable Y .

La valeur par défaut du nombre de retards est généralement fixée à 2 dans plusieurs logiciels statistiques. Mentionnons enfin que la stationnarité des séries est une exigence majeure pour la validité des résultats. En pratique, une transformation des séries (logarithme, différences et/ou différences saisonnières, etc.) est souvent nécessaire avant de procéder à l'analyse de causalité.

Pour satisfaire la condition de stationnarité, les séries ont été préalablement transformées par le logarithme suivi d'une différence première et d'une différence saisonnière. Dans un premier temps nous considérons les séries dans toute leur étendue, c'est à dire sur la période allant de janvier 1981 à avril 2003. Ensuite, nous considérons les mêmes séries pour les sous périodes allant respectivement de janvier 1981 à décembre 1997 et de janvier 1998 à avril 2003. Cette subdivision est quelque peu arbitraire, et son choix est simplement motivé par le fait qu'elle semble s'accorder (au moins graphiquement) au comportement de l'IOE par rapport aux séries cibles.

Rappelons à ce niveau que l'hypothèse nulle dans le test de Granger est: X ne cause pas Y à la Granger.

La table 6.3 présente les seuils observés du test de Granger pour la période complète (janvier 1981 - avril 2003). Dans le cas du taux d'emploi et l'IOE par exemple, les seuils observés sont respectivement de 3.2E-6 et 1.31E-12. Ce qui signifie qu'en prenant un seuil nominal α de 1% on rejeterait non seulement la non causalité du taux d'emploi sur l'IOE, mais aussi la non causalité de l'IOE sur le taux d'emploi. Ce résultat suggérerait une causalité bi-directionnelle entre l'IOE et le taux d'emploi, ce qui semble contraire au bon sens. Ce phénomène est courant dans l'analyse de causalité grangerienne et est connu sous le vocable *problème de feedback*. Dans de telles situations, les résultats du test F n'ont plus la même signification. Il serait fort probable dans ce cas-ci qu'une ou plusieurs autres variables causent simultanément l'IOE et le

taux d'emploi. Ainsi, la relation existante entre les deux variables est une relation induite par l'effet de cette (ou de ces) autre variable exogène.

Considérons maintenant les résultats de la sous période allant de janvier 1981 à décembre 1997 (voir table 6.4). Globalement, les seuils observés, quoique faibles, sont plus élevés que ceux observés dans le cas précédent. Cependant, ici aussi, l'on observe le même problème de feedback entre l'IOE et les autres variables.

Dans le cas de la sous période allant de janvier 1998 à avril 2003 (voir table 6.5), les résultats sont différents de ceux des cas précédents. En effet, les seuils observés sont plus grands et montrent que l'hypothèse de non causalité des variables de l'emploi vers l'IOE, est fortement acceptée. Ainsi, dans le cas du taux d'emploi et l'IOE par exemple, les seuils observés sont respectivement de 0.3024 et 6.49E-6, par conséquent, seule la causalité de l'IOE sur le taux d'emploi est significative. Cependant, il faut être prudent dans son interprétation étant donné le feedback observé entre l'IOE et les séries cibles dans les autres périodes.

Pour conclure cette sous section, mentionnons que les résultats du test de Granger sont fortement contaminés par le problème de feedback entre l'IOE et les autres variables. Il ressort de cette analyse que la relation qui semble exister entre l'IOE et ces variables n'est pas nécessairement de nature causale. Cette relation est probablement induite par l'effet d'une (ou d'autres) variable externe qui cause simultanément l'IOE et les variables cibles.

4. MODÉLISATION ET PRÉVISION

Dans les sections précédentes, nous avons montré l'évidence d'une relation (non nécessairement causale) entre l'IOE et les variables de l'emploi. La suite logique de cette analyse serait donc de trouver la forme de cette relation dans le cadre de l'évaluation du pouvoir de prévision de l'IOE. Dans cette démarche, nous ignorerons momentanément le problème de feedback rencontré dans l'analyse de causalité pour ajuster des modèles de fonction de transfert (ou encore modèles de Box-Tiao, etc.) que nous comparerons aux modèles univariés obtenus pour chacune des séries cibles. Les modèles de fonction de transfert des modèles ARIMA généralisés avec présence de variables exogènes.

Dans cette section, nous ajustons d'abord un modèle ARIMA univarié et un modèle de fonctions de transfert pour chacune des séries de l'emploi pour la période allant de janvier 1981 à avril 2003. Ensuite, à l'aide de ces modèles, nous calculons les prévisions de la série cible, un mois à la fois, pour la période allant de janvier 1998 à avril 2003. Par la suite, nous comparons les erreurs de prévision pour les deux types de modèles.

Pour toutes les séries, le logarithme, la première différence et la différence saisonnière ont été utilisés pour avoir un modèle ARIMA satisfaisant (paramètres significatifs et résidus non corrélés). Dans le cas du taux d'emploi par exemple, le modèle univarié s'écrit:

$$(1 - 0.173B^3) \nabla \nabla_{12} \text{Log}(TE_t) = (1 + 0.337B^2 + 0.16B^4)(1 - 0.53B^{12}) \epsilon_t$$

où B , désigne l'opérateur de retard. Le modèle de fonctions de transfert correspondant s'écrit:

$$\nabla \nabla_{12} \text{Log}(TE_t) = (0.012 + 0.02B + 0.02B^2) \nabla \nabla_{12} \text{Log}(IOE_t) + (1 + 0.17B^2)(1 - 0.547B^{12}) \epsilon_t$$

La table 6.5 présente les seuils observés du test t d'égalité des moyennes des erreurs de prévision pour les deux types de modèles. Aussi, elle présente quelques statistiques usuelles (moyenne, médiane et écart type) sur ces erreurs de prévision. Il ressort que pour chacune des séries cibles, on ne peut rejeter l'hypothèse d'égalité des moyennes de prévision (les seuils observés sont tous supérieurs à 0.05). Ainsi, la prévision des variables de l'emploi à partir de leur historique conjoint avec l'IOE n'est pas significativement meilleure à celle obtenue à partir de leur seul passé.

5. CONCLUSION

Dans ce travail, nous avons mené, à travers trois types d'analyse, une étude sur le pouvoir de prévision de l'indice d'offre d'emploi par rapport à quelques variables pertinentes de l'emploi. L'analyse graphique et l'analyse de causalité ont montré l'évidence d'une relation, non nécessairement causale, de l'IOE vers ces variables. Ces analyses ont aussi montré que l'intensité de la relation diminue dans le temps. Cependant, à travers l'analyse de modélisation et de prévision, il ressort qu'il est difficile d'obtenir une forme analytique de cette relation qui permettrait de mieux prédire les variables de l'emploi à partir de l'IOE. Les modèles de fonctions de transfert obtenus, quoique satisfaisants, ne performant pas mieux que les modèles univariés. Il faudrait donc essayer d'améliorer autant que possible ces modèles multi variés.

En conclusion, nous pouvons dire qu'à travers cette étude, il n'y a pas suffisamment d'évidence statistique qui supporte l'existence d'un pouvoir de prévision significatif de l'IOE sur les variables de l'emploi et du chômage.

BIBLIOGRAPHIE

Akeyampong, Ernest B. et Lacroix, Justin (2002). The Internet and Statistics Canada's Help Wanted Index (HWI). Staff reports, Labor and Household Surveys Analysis Division, Statistics Canada No. 02-2002E.

Amoah, Benjamin (2000). Help Wanted Index. Perspectives on Labour and Income, **12**, No. 2, Catalogue 75-001 (Statistics Canada).

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods'. *Econometrica*, **37**, 424-438.

Haggard-Guénette, Cynthia (1989). Job Ads: A Leading Indicator?. *Perspectives on Labor and Income*, **1**, No. 2 53-63 (Statistics Canada).

Sims, C. (1972). "Money, income and causality". *American Economic Review*, **62**, 540-542.

APPENDICE

Figure 6.1 - graphique des séries lissées (v-3) avec la série lissée (v-25) de l'IOE.

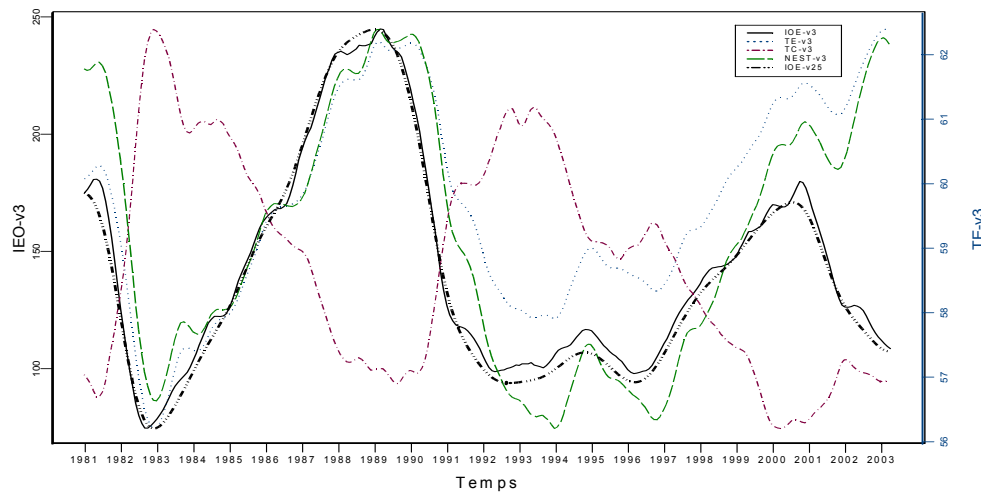


Figure 6.2 : - graphique de toutes les séries lissées avec le filtre v-25.

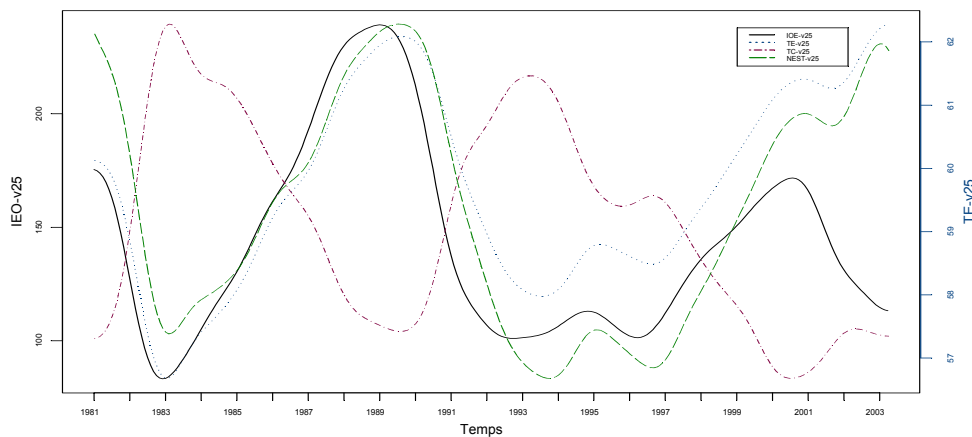


Table 6.1 - Mauvaises prédictions de l'IOE pour le filtre v-3 : (+) un retournement se produit dans la série cible sans que IOE ne l'annonce; (-) IOE annonce un retournement qui ne se produit pas dans la série cible

TC	TE	NEST
		S-Oct83(+)
C-Dec83(+)	C-Jul89(+)	C-Mar84(+)
	S-Fev90(+)	C-Jul89(+)
C-Mai92(-)	C-Mai92(-)	S-Fev90(+)
	S-Mai93(-)	C-Mai92(-)
C-Sep93(-)	C-Sep93(-)	S-Mai93(-)
S-Nov94(-)	C-Sep95(+)	C-Jan94(+)
C-Nov95(+)	S-Fev96(+)	
C-Avr2000(+)	S-Avr2000(+)	C-Nov2001(+)
S-Oct2000(-)	C-Nov2001(+)	
7MP	9MP	8MP

Table 6.2 - Points de retournements pour le filtre v-25

IOE	TC	TE	NEST
C-Dec82	S-Fev83 (+2)	C-Jan83 (+1)	C-Fev83 (+2)
S-Jan89	C-Aou89 (+7)	S-Sep89 (+8)	S-Aou89 (+7)
C-Oct92	S-Mar93 (+5)		
		C-Dec93 (+14)	C-Mar94 (+17)
S-Nov94		S-Mar95 (+4)	S-Fev95 (+3)
	C-Nov95 (+12)		
C-Avril96	S-Sep96 (+5)	C-Sep96 (+5)	C-Sep96 (+5)
S-Aou2000	C-Jul2000 (-1)	S-Nov2000 (+3)	S-Dec2000 (+4)
		C-Sep2001	C-Sep2001
	S-Mai2002		
6PR (OFS)	7PR (OFS)	7PR (OFS)	7PR (OFS)

Table 6.3 - Seuils observés pour le test F de Granger pour la période complète

X	Y			
	IOE	TC	TE	NE
IOE		2.52E-8	1.31E-12	6.25E-12
TC	4.76E-5			
TE	3.2E-6			
NE	100.32E-6			

Table 6.4 - Seuils observés pour le test F de Granger pour la période : janvier 1981 – décembre 1997

<i>X</i>	<i>Y</i>			
	IOE	TC	TE	NE
IOE		1.95E-6	5.65E-9	1.28E-8
TC	2.12E-4			
TE	1.59E-6			
NE	1.64E-6			

Table 6.5 : Seuils observés pour le test F de Granger pour la période : janvier 1998 – avril 2003

<i>X</i>	<i>Y</i>			
	IOE	TC	TE	NE
IOE		0.01863	6.49E-6	1.04E-5
TC	0.3514			
TE	0.3024			
NE	0.095			

Table 6.6 - Moyenne, Médiane et Écart type des erreurs relatives de prévision et seuils observés du test t.

Séries	Modèles Univariés	Modèles de FT	Seuils observés du test t d'égalité des moyennes
TE	0.002503	0.0021	0.1127
	0.002256	0.00184	
	0.00151	0.00133	
TC	0.021945	0.021309	0.8435
	0.015424	0.015607	
	0.01921	0.01709	
NE	0.00215	0.001904	0.3363
	0.002082	0.00145	
	0.00149	0.0014	

